

SELF-INTERSECTIONS OF RANDOM GEODESICS ON NEGATIVELY CURVED SURFACES

STEVEN P. LALLEY

ABSTRACT. We study the fluctuations of self-intersection counts of random geodesic segments of length t on a compact, negatively curved surface in the limit of large t . If the initial direction vector of the geodesic is chosen according to the *Liouville measure*, then it is not difficult to show that the number $N(t)$ of self-intersections by time t grows like κt^2 , where $\kappa = \kappa_M$ is a positive constant depending on the surface M . We show that (for a smooth modification of $N(t)$) the fluctuations are of size t , and the limit distribution is a weak limit of Gaussian quadratic forms. We also show that the fluctuations of *localized* self-intersection counts (that is, only self-intersections in a fixed subset of M are counted) are typically of size $t^{3/2}$, and the limit distribution is Gaussian.

1. FLUCTUATIONS OF SELF-INTERSECTION COUNTS

Choose a point x and a direction θ at random on a compact, negatively curved surface M , and let $\gamma(t) = \gamma(t; x, \theta)$ be the (unit speed) geodesic ray in direction θ started at x . For large t the number $N(t) = N(\gamma[0, t])$ of transversal¹ self-intersections of the geodesic segment $\gamma[0, t]$ will be of order t^2 ; in fact, if $\kappa_M = (2\pi|M|)^{-1}$, where $|M|$ denotes the surface area of M , then

$$(1) \quad \lim_{t \rightarrow \infty} N(t)/t^2 = \kappa_M/2$$

with probability 1. See section 3 below for the (easy) proof. Furthermore, the empirical distribution of the self-intersection points converges to the uniform distribution on the surface. Similar results hold for a randomly chosen *closed* geodesic [8]: If from among all closed geodesics of length $\leq L$ one is chosen at random, then the number of self-intersections, normalized by L^2 , will, with probability approaching one as $L \rightarrow \infty$, be close to κ_g . These results have been extended [12] to the number and distribution of self-intersections at angles in fixed intervals $[\alpha, \beta]$. Closed geodesics with *no* self-intersections have long been of interest in geometry — see, for instance, [3, 2] — and recently, M. Mirzakhani [10] found the asymptotic growth rate of the number of simple closed geodesics of length $\leq t$ as $t \rightarrow \infty$. That this number is not 0 shows (in view of the Law of Large Numbers (1)) that there is substantial variation in the random variable $N(t)$.

The primary objective of this paper is to investigate the *fluctuations* (second-order asymptotics) of the self-intersection numbers. One's first guess might be that these are of order t , and this is indeed the case; however, lest this seem too obvious we add that if one counts only self-intersections

Date: July 1, 2009.

1991 *Mathematics Subject Classification*. Primary 58F17, secondary 53C22, 37D20.

Key words and phrases. geodesic flow, self-intersection, central limit theorem, multiple mixing.

Supported by NSF grant DMS-0405102.

¹If the initial point x and direction θ are chosen randomly (according to the normalized Liouville measure on the unit tangent bundle) then there is probability 0 that the resulting geodesic will be periodic, so with probability 1 every self-intersection will necessarily be transversal.

in a (nice) sub-domain $U \subset M$ then the fluctuations are no longer of order t , but rather $t^{3/2}$. (In fact we will only prove these statements for smoothed versions of the counts.) One might also guess that the rescaled random variable $(N(t) - \kappa_g t^2)/t$ should converge in distribution as $t \rightarrow \infty$ to a Gaussian distribution, but this, as we will show, is probably false: the limit distribution is a weak limit of *Gaussian quadratic forms*. (This does not preclude the possibility that it is Gaussian, but the arguments below will make it clear that this is unlikely.) Weak limits of Gaussian quadratic forms are known to occur in connection with stationary processes exhibiting long-range dependence [16], [9] (where they are known as *Rosenblatt distributions*), and also in the connection with certain types of U -statistics [15].

Definition 1. A *Gaussian quadratic form* is a random variable (or its distribution) of the form $\sum_{j=1}^m \theta_j Z_j^2$, where the random variables Z_j are independent, unit Gaussians and θ_j are real scalars.

Unfortunately, the study of fluctuations in the self-intersection numbers $N(t)$ is complicated by the (infrequent) occurrence of self-intersections at very small angles. We have not yet been able to successfully resolve the technical issues created by such self-intersections, and so we will state our main result not for the counts $N(t)$ but rather for a smoothed version $N_\varphi(t)$ defined as follows. Let $\varphi : \mathbb{R} \rightarrow [0, \infty)$ be an *even*, C^∞ , nonnegative, 2π -periodic function. For each self-intersection i of the geodesic segment $\gamma[0, t]$, denote by $\theta_i \in [0, 2\pi)$ the angle of the self-intersection: in particular, if the self-intersection occurs at times $0 \leq s_i < t_i \leq t$, then θ_i is the angle between the tangent vector to γ at s_i and the tangent vector at t_i . Define the *smoothed self-intersection number*

$$(2) \quad N_\varphi(t) = N_\varphi(\gamma[0, t]) = \sum_{i=1}^{N(t)} \varphi(\theta_i).$$

Clearly, if $\varphi \equiv 1$ then $N_\varphi(t) = N(t)$. We will prove in section 3 that the smoothed self-intersection numbers satisfy a strong law of large numbers (SLLN) analogous to (1):

$$(3) \quad \lim_{t \rightarrow \infty} N_\varphi(t)/t^2 = \kappa_\varphi/2$$

where

$$(4) \quad \kappa_\varphi = \frac{1}{2\pi|M|} \int_0^{2\pi} \varphi(\theta) |\sin \theta| d\theta.$$

Theorem 1. Assume that the smoothing function φ is zero in a neighborhood $[-\alpha, \alpha]$ of 0 and positive in $(\alpha, \pi - \alpha)$. If the initial point x and direction θ are chosen randomly according to the Liouville measure, then as $t \rightarrow \infty$,

$$(5) \quad \frac{N_\varphi(t) - \kappa_\varphi t^2}{t} \implies \Psi$$

where \implies indicates weak convergence to a distribution Ψ in the weak closure of the set of Gaussian quadratic forms. This limiting distribution may depend on both the surface M and the smoothing function φ .

We have not been able to prove that the limit distribution Ψ is nondegenerate, nor that it is non-Gaussian, but this seems unlikely (see Section 4.1 below). It is naturally of interest to consider also fluctuations in the empirical distribution of self-intersection points. Let $f : M \rightarrow \mathbb{R}$ be any continuous function on the surface M . For each self-intersection i of the geodesic segment $\gamma[0, t]$, denote by $\theta_i \in [-\pi, \pi]$ the angle and $x_i \in M$ the location of the self-intersection. Define the

f -localized self-intersection counts

$$(6) \quad N_{\varphi;f}(t) = N_{\varphi;f}(\gamma[0, t]) = \sum_{i=1}^{N(t)} f(x_i) \varphi(\theta_i).$$

As for the global self-intersection counts, the localized self-intersection random variables $N_{\varphi;f}(t)$ obey a strong law of large numbers: For ν_L -almost every initial direction vector,

$$(7) \quad \lim_{t \rightarrow \infty} N_{\varphi;f}(t)/t^2 = \frac{\kappa_\varphi}{2} \int_M f(x) dx := A_{\varphi;f},$$

where dx indicates the normalized surface area measure on the surface M . See section 5.1 for the proof. What is interesting about the localized self-intersection counts is that their fluctuations are of a different order of magnitude than those of the global variables $N_\varphi(t)$, at least for functions f of small support:

Theorem 2. *For any compact, negatively curved surface M there is a constant $\varepsilon > 0$ such that the following is true: If the smoothing function φ satisfies the hypotheses of Theorem 1; if $f : M \rightarrow \mathbb{R}$ is any C^∞ , nonnegative function that is not identically 0 and whose support has diameter $< \varepsilon$; and if the initial direction vector $\gamma(0)$ is chosen randomly according to the Liouville measure, then for some $\sigma = \sigma_{\varphi;f} > 0$,*

$$(8) \quad \frac{N_{\varphi;f}(t) - A_{\varphi;f}t^2}{\sigma t^{3/2}} \Longrightarrow \Phi$$

as $t \rightarrow \infty$, where Φ is the standard (mean 0, variance 1) Gaussian distribution.

Simple lower bounds for $\varepsilon > 0$ will be given in Section 5, where Theorem 2 will be proved. Theorem 1 will be proved in section 4, using an extension of a mixing result of Dolgopyat [5]. The proof relies on a representation of the self-intersection counts in terms of what we dub the *intersection kernels* for the geodesic flow; these are studied in section 2. For completeness, we present proofs of the strong laws of large numbers (1) and (3) in section 3, and of (7) in section 5.1.

2. INTERSECTION KERNEL

2.1. The intersection kernel. Geodesics on any surface², regardless of its curvature, look locally like straight lines. Consequently, for any compact surface M with smooth Riemannian metric there exists $\varrho > 0$ such that if α and β are geodesic segments of length $\leq \varrho$ then α and β intersect transversally, if at all, in at most one point. Thus, if γ is not periodic then the smoothed self-intersection number $N_\varphi(L) = N_\varphi(\gamma[0, L])$ can be computed by partitioning $\gamma[0, L]$ into nonoverlapping segments of common length $\delta \leq \varrho$ and counting the number of pairs that intersect transversally. Let α_i and α_j be two such segments; then the event that these segments intersect is completely determined by their initial points and directions, as is the angle of intersection.

Definition 2. The *intersection kernel* $H_\delta : SM \times SM \rightarrow \mathbb{R}_+$ is the nonnegative function that takes the value $H_\delta(u, v) = \varphi(\theta)$ if the geodesic segments of length δ with initial tangent vectors u and v intersect transversally at angle θ , and $H_\delta(u, v) = 0$ otherwise.

²Here and throughout the paper the term *geodesic* will be used either to indicate a geodesic path in the surface M or its lift to the unit tangent bundle SM ; the meaning should be clear from context. Two geodesic segments will be said to intersect if their projections to the surface M intersect.

The dependence on the smoothing function φ is suppressed, as φ will be fixed throughout the paper. Because φ is assumed to be even and π -periodic, H_δ is symmetric in its arguments u, v . The intersection kernel determines the smoothed self-intersection numbers as follows: If $L = n\delta$ is an integer multiple of δ , then for any geodesic γ ,

$$(9) \quad N_\varphi(L) = N_\varphi(\gamma[0, L]) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n H_\delta(\gamma(i\delta), \gamma(j\delta)).$$

The factor of $1/2$ compensates for the double-counting that results from letting both indices of summation i, j range over all n geodesic segments. Note that the diagonal terms $H_\delta(\gamma(i\delta), \gamma(i\delta))$ in this sum are all 0, because the segment $\gamma(i\delta)$ does not intersect itself transversally.

2.2. The associated integral operators. The intersection kernel $H_\delta(u, v)$ is symmetric in its arguments and Borel measurable, but not continuous, because self-intersections can be created or destroyed by small perturbations of the initial vectors u, v . Nevertheless, H_δ induces a self-adjoint integral operator on the Hilbert space $L^2(\nu_L)$ by

$$(10) \quad H_\delta \psi(u) = \int_{v \in SM} H_\delta(u, v) \psi(v) d\nu_L(v).$$

Lemma 1. *For all sufficiently small $\delta > 0$,*

$$(11) \quad H_\delta 1(u) := \int H_\delta(u, v) d\nu_L(v) = \delta^2 \kappa_\varphi$$

for all $u \in M$. Thus, the constant function 1 is an eigenfunction of the operator H_δ , and consequently the normalized kernel $H_\delta(u, v)/\delta^2 \kappa_\varphi$ is a Markov kernel.

Remark 1. The integral $H_\delta 1(u)$ is the expectation of $\varphi(\theta)$ where θ is the angle in which a randomly chosen geodesic segment of length δ intersects the geodesic segment of length δ with initial tangent vector u . The assertion of the lemma is that this expectation does not depend on the initial tangent vector u , even for surfaces M of variable negative curvature.

Proof. Denote by $\alpha = \gamma([0, \delta]; u)$ the geodesic segment of length δ with initial tangent vector u . For small $\delta > 0$ and fixed angle θ , the set of points $x \in S$ such that a geodesic segment of length δ with initial base point x intersects α at angle θ is approximately a rhombus of side δ with an interior angle θ . The area of such a rhombus is $\delta^2 |\sin \theta|$. Hence, as $\delta \rightarrow 0$,

$$(12) \quad \int H_\delta(u, v) d\nu_L(v) \sim \delta^2 \int_0^{2\pi} \varphi(\theta) |\sin \theta| d\theta / (2\pi |M|) = \delta^2 \kappa_\varphi,$$

and the relation \sim holds uniformly for $u \in SM$.

It remains to show that the approximate equality \sim is actually an equality for small $\delta > 0$, equivalently, that the value of the integral stabilizes as $\delta \rightarrow 0$. Assume that $\delta > 0$ is sufficiently small that any two distinct geodesic segments of length δ intersect transversally at most once. Consider the geodesic segments of length δ with initial direction vectors u and v . For any integer $m \geq 2$, each of these segments can be partitioned into m nonoverlapping sub-segments (each open on one end and closed on the other) of length δ/m . At most one pair of these constituent sub-segments can intersect; hence,

$$H_\delta(u, v) = \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} H_{\delta/m}(\gamma(i\delta; u), \gamma(j\delta; v)).$$

Integrating over v with respect to the Liouville measure ν_L and using the invariance of ν_L relative to the geodesic flow we obtain that

$$H_\delta 1(u) = \sum_{i=0}^{m-1} m H_{\delta/m} 1(\gamma(i\delta; u)).$$

Let $m \rightarrow \infty$ and use the approximation (12) (with δ replaced by δ/m); since this approximation holds uniformly, it follows that $H_\delta 1(u) = \delta^2 \kappa_S$. \square

Lemma 2. *For each $\delta > 0$ sufficiently small, the integral operator H_δ on $L^2(\nu_L)$ is compact.*

Proof. If the kernel $H_\delta(u, v)$ were jointly continuous in its arguments u, v then this would follow by standard results about integral operators — see, e.g., [17]. Since H_δ is not continuous, these standard results do not apply; nevertheless, the argument for compactness is elementary. It suffices to show that the mapping $u \mapsto H_\delta(u, \cdot)$ is continuous relative to the L^2 -norm. Take $u, u' \in SM$, and let α, α' be the geodesic segments of length δ started at u, u' , respectively. If u, u' are close, then the geodesic segments α, α' are also close. Hence, for all but very small angles θ the set of points $x \in M$ such that a geodesic segment of length δ with initial base point x intersects α at angle θ but does not intersect α' is small. Consequently, the functions $H_\delta(u, \cdot)$ and $H_\delta(u', \cdot)$ differ on a set of small measure. \square

Lemma 2 implies that the Hilbert-Schmidt theory applies. In particular, the non-zero spectrum of H_δ consists of isolated real eigenvalues λ_j of finite multiplicity (and listed according to multiplicity). The corresponding (real) eigenfunctions ψ_j can be chosen so as to constitute an orthonormal basis of $L^2(\nu_L)$, and the eigenvalue sequence λ_j is square-summable.

Lemma 3. *The Markov kernel $\bar{H}_\delta := H_\delta / \delta^2 \kappa_\varphi$ satisfies the Doeblin condition: There exist an integer $n \geq 1$ and a positive real number ε such that*

$$(13) \quad \bar{H}_\delta^n(u, v) \geq \varepsilon \quad \text{for all } u, v \in SM.$$

Proof. Chose n so large that for any two points $x, y \in S$ there is a sequence $\{x_i\}_{0 \leq i \leq n}$ of $n+1$ points beginning with $x_0 = x$ and ending at $x_n = y$, and such that each successive pair x_i, x_{i+1} are at distance $< \delta/4$. Then for any two geodesic segments α, β of length δ on S there is a chain of $n+1$ geodesic segments α_i , all of length δ , beginning at $\alpha_0 = \alpha$ and ending at $\alpha_n = \beta$, such that any two successive segments α_i and α_{i+1} intersect transversally. Since the intersections are transversal, the initial points and directions of these segments can be jiggled slightly without undoing any of the transversal intersections. This implies (13). \square

Corollary 1. *The eigenvalue $\delta^2 \kappa_\varphi$ is a simple eigenvalue of the integral operator H_δ , and the rest of the spectrum lies in a disk of radius $< \delta^2 \kappa_\varphi$.*

Proof. This is a standard result in the theory of Markov operators. \square

Corollary 2. *For every $j \geq 2$ the eigenfunction ψ_j has mean zero relative to ν_L , and distinct eigenfunctions are uncorrelated.*

Proof. The spectral theorem guarantees orthogonality of the eigenfunctions. The key point is that $\psi_1 = 1$ is an eigenfunction, and so the orthogonality $\psi_j \perp \psi_1$ implies that each ψ_j for $j \geq 2$ has mean zero. \square

Lemma 4. *If $\delta > 0$ is sufficiently small then H_δ has eigenvalues other than 0 and $\lambda_1(\delta)$.*

Proof. Otherwise, the Markov operator \bar{H}_δ would be a projection operator: for every $\psi \in L^2(\nu_L)$ the function $\bar{H}_\delta\psi$ would be constant. But if $\delta > 0$ is small, this is obviously not the case. \square

2.3. Smoothing. The discontinuity of the kernel H_δ creates certain technical problems: for instance, the eigenfunctions ψ_j need not be continuous. Thus, it will be to our advantage to approximate H_δ by a smooth kernel K_δ in such a way that the sums (9) are not too badly disturbed when H_δ is replaced by the approximation K_δ . It is *solely* for the purpose of constructing this approximation that the restrictions in Theorem 1 on the smoothing kernel φ — in particular, that it vanishes in a neighborhood of 0 — are needed.

Let $p(s)$ be an even, C^∞ probability density on \mathbb{R} with support contained in the interval $[-1, 1]$. Define \mathcal{D} to be the set of pairs $(u, v) \in SM \times SM$ such that the geodesics through u and v cross, if at all, *transversally*; this set is dense in $SM \times SM$. For any pair $(u, v) \in \mathcal{D}$, the (two-sided) geodesics $\{\gamma(s; u)\}_{s \geq 0}$ and $\{\gamma(t; v)\}_{t \geq 0}$ will intersect in at most countably many points, which can be labeled (s_i, t_i) where the entries s_i and t_i denote the signed distances along the two geodesics from their origins u, v where the intersection occurs. (The intersections do not generally occur in the same (time) order along the two geodesics.) Let $\theta_i \in (0, \pi)$ be the angle of crossing at the i th intersection, and define

$$(14) \quad K_\delta(u, v) = \sum_i \delta^{-2} p(s_i/\delta) p(t_i/\delta) \varphi(\theta_i).$$

Lemma 5. *Assume that the smoothing function φ is zero in a neighborhood $[-\alpha, \alpha]$ of 0. If $\delta > 0$ is sufficiently small then for any pair $(u, v) \in \mathcal{D}$, the sum (14) contains at most one nonzero term, so $K_\delta(u, v)$ is well-defined and finite. Furthermore, the function $K_\delta(u, v)$ extends to a C^∞ , symmetric function on $SM \times SM$, by setting $K_\delta(u, v) = 0$ for all $(u, v) \notin \mathcal{D}$.*

Proof. Since φ is even, the kernel K_δ is symmetric if the sum (14) is finite. Now in order that the i th term of the sum (14) be nonzero, the distances $|s_i|$ and $|t_i|$ must both be smaller than δ , because p has support contained in $[-1, 1]$. But if $\delta > 0$ is sufficiently small then any two geodesic segments of length 2δ will intersect transversally at most once. Thus, the sum (14) has at most one nonzero term.

Since geodesics vary smoothly with their initial conditions, the intersection distances s_i, t_i and angles θ_i vary smoothly with u, v in \mathcal{D} . Consequently, the function K_δ is C^∞ in \mathcal{D} . But as (u, v) approaches the boundary of \mathcal{D} , the angle(s) of intersection of the geodesics through u and v must approach zero. Hence, by the assumption on φ , the kernel K_δ vanishes near $\partial\mathcal{D}$. \square

Lemma 6. *For any geodesic γ , the smoothed self-intersection number $N_\varphi(t)$ of the segment $\gamma[0, t]$ satisfies the inequalities*

$$(15) \quad \frac{1}{2} \iint_{[\delta, t-\delta]^2} K_\delta(\gamma(r), \gamma(s)) dr ds \leq N_\varphi(t) \leq \frac{1}{2} \iint_{[-\delta, t+\delta]^2} K_\delta(\gamma(r), \gamma(s)) dr ds.$$

Proof. Suppose that γ has a self-intersection at some (r, s) , that is, the vectors $\gamma(r)$ and $\gamma(s)$ lie over the same base point in M . Then by definition of K_δ , since h is a probability density supported by $[-\delta/2, \delta/2]$, the integral of $K_\delta(\gamma(r'), \gamma(s'))$ over the square of side δ centered at (r, s) must be 1. Hence, both bounding integrals in (15) count each such self-intersection (r, s) with weight 1. The only self-intersections not counted correctly are those (r, s) where either r or s lies within δ of

one of the time endpoints 0 or t . Adjusting the limits of integration by $\pm\delta$ compensates for these boundary errors. \square

Remark 2. The inequalities (15) imply that the double integral on the right side of (15) is bounded above by $N_\varphi(\gamma[-\delta, t + \delta])$, and the double integral on the left is bounded below by $N_\varphi(\gamma[\delta, t - \delta])$. Therefore, the errors in the inequalities (15) are no larger than

$$(16) \quad N_\varphi(\gamma[-\delta, t + \delta]) - N_\varphi(\gamma[\delta, t - \delta]).$$

\square

The smoothed kernels $K_\delta(u, v)$ enjoy all of the properties enumerated for the intersection kernels $H_\delta(u, v)$ in sec. 2.2 above, provided the smoothing window $\delta > 0$ is sufficiently small. In particular,

- (P1) The constant function 1 is an eigenfunction of K_δ , with eigenvalue κ_φ .
- (P2) The integral operator K_δ is compact.
- (P3) The Markov kernel K_δ/κ_φ satisfies the Doeblin condition.
- (P4) The eigenvalue κ_φ is simple, and the rest of the spectrum lies in some $[-\kappa_\varphi + \varepsilon, \kappa_\varphi - \varepsilon]$.
- (P5) Distinct eigenfunctions are uncorrelated, and except for the constant eigenfunction have mean 0.

These may all be proved by mimicking the proofs of the analogous assertions for the self-intersection kernels H_δ . In the special case that the smoothing density p in the definition (14) is not only even but also nondecreasing on $(-\infty, 0]$, property (P1) can be deduced directly from Lemma 1, because in this case the kernel K_δ is a convex combination of the kernels $H_\varepsilon/\varepsilon^2$, where $0 < \varepsilon \leq \delta$. Only Property (P1) will be needed for the proof of Theorem 1, and we will have no need to consider smoothing densities p that are not monotone on $(-\infty, 0]$, so we refrain from spelling out the details of the proofs of (P1)—(P5) in the general case.

3. SLLN FOR SELF-INTERSECTIONS

3.1. SLLN for the smoothed self-intersection numbers. According to Lemmas 5–6, if the smoothing function φ is zero in a neighborhood $[-\alpha, \alpha]$ of 0, then the intersection kernel H_δ can be approximated by a continuous kernel K_δ in such a way that the smoothed self-intersection number $N_\varphi(t)$ is well-approximated by

$$N^*(\gamma[0, t]) := \frac{1}{2} \iint_{[0, t]^2} K_\delta(\gamma(r), \gamma(s)) dr ds.$$

For the strong law of large numbers (3), only a crude bound on the error in this approximation is needed. By Remark 2, the error is bounded by (16); and by another use of the double inequalities (15), it follows that the discrepancy is bounded by the difference $N^*(\gamma[-\delta, t + \delta]) - N^*(\gamma[\delta, t - \delta])$. Consequently, since $\delta > 0$ can be chosen arbitrarily small, to prove the strong law of large numbers (3) it suffices to show that for $\delta > 0$ sufficiently small and for ν_L -almost every initial point $\gamma(0)$,

$$(17) \quad \lim_{t \rightarrow \infty} t^{-2} \iint_{[0, t]^2} K_\delta(\gamma(s_1), \gamma(s_2)) ds_1 ds_2 = \kappa_\varphi$$

The following proposition implies that for any probability measure μ on SM that is invariant and ergodic under the geodesic flow, for μ -almost every initial point $\gamma(0)$,

$$(18) \quad \lim_{t \rightarrow \infty} t^{-2} \iint_{[0,t]^2} K_\delta(\gamma(s_1), \gamma(s_2)) ds_1 ds_2 = \iint_{SM \times SM} K_\delta(x, y) d\mu(x) d\mu(y).$$

That the expectation on the right equals κ_φ for $\mu = \nu_L$ follows from property (P1) above.

Proposition 1. *Let (\mathcal{X}, d) be a compact metric space and let $K : \mathcal{X}^2 \rightarrow \mathbb{R}$ be continuous. If μ is a Borel probability measure on \mathcal{X} and $T : \mathcal{X} \rightarrow \mathcal{X}$ is an ergodic, measure-preserving transformation (not necessarily continuous) relative to μ , then*

$$(19) \quad \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n K(T^i x, T^j x) = \iint_{\mathcal{X} \times \mathcal{X}} K(y, z) d\mu(y) d\mu(z)$$

for μ -almost every x .

Proof. The function K is bounded, since it is continuous, so the double integral in (19) is well-defined and finite. Furthermore, the set of functions K_x defined by $K_x(y) := K(x, y)$, where x ranges over the space \mathcal{X} , is equicontinuous, and the function

$$\bar{K}_x := \int_{\mathcal{X}} K_x(y) d\mu(y)$$

is continuous in x . The equicontinuity of the functions K_x implies, by the Arzela-Ascoli theorem, that for any $\varepsilon > 0$ there is a finite subset $F_\varepsilon = \{x_i\}_{1 \leq i \leq I}$ such that for any $x \in SM$ there is at least one index $i \leq I$ such that

$$\|K_x - K_{x_i}\|_\infty < \varepsilon.$$

It follows that the time average of K_x along any trajectory differs from the corresponding time average of K_{x_i} by less than ε . But since the set F_ε is finite, Birkhoff's ergodic theorem implies that if the initial point and direction of γ are chosen randomly according to ν_L then with probability one, for each $x_i \in F_\varepsilon$,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t K(x_i, \gamma(s)) ds = \int K(x_i, y) d\nu_L(y).$$

It therefore follows from equicontinuity (let $\varepsilon \rightarrow 0$) and the continuity in x of the averages \bar{K}_x that almost surely

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t K(x, \gamma(s)) ds = \int K(x, y) d\nu_L(y)$$

uniformly for $x \in \mathcal{X}$. The uniformity of this convergence guarantees that (19) holds μ -almost surely. \square

Remark 3. Wiener's multi-parameter ergodic theorem ([18], Theorems I'' – II'') implies that under much weaker hypotheses³ on the function K ,

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n K(T^i x, T^j y) = \iint_{\mathcal{X} \times \mathcal{X}} K(u, v) d\mu(u) d\mu(v) \quad \text{for } (\mu \times \mu)\text{-every } (x, y).$$

Proposition 1 cannot be deduced from this, as the diagonal of $\mathcal{X} \times \mathcal{X}$ has $(\mu \times \mu)$ -measure 0. In fact, it is *not* generally true that the convergence (19) holds for functions $K(x, y)$ that are not

³Wiener requires that $K \in L^2(\mu \times \mu)$. For still weaker hypotheses, see [6].

continuous, even if the measure-preserving transformation T is mixing. A simple example can be constructed as follows. Let $R = R_\theta$ be an irrational rotation of the circle S^1 , and let $\sigma : \Sigma \rightarrow \Sigma$ be the shift on the space of all one-sided sequences $\omega = \omega_1\omega_2\cdots$ with entries ± 1 . For $x \in S^1$ and $\omega \in \Sigma$, define

$$T(x, \omega) = (R^{\omega_1}x, \sigma\omega);$$

this is a mixing, measure-preserving transformation relative to $\lambda = \text{Lebesgue} \times \mu$, where μ is the product Bernoulli-(1/2) measure on Σ . Now for $x, y \in S^1$ and $\omega, \omega' \in \Sigma$, define

$$\begin{aligned} K((x, \omega), (y, \omega')) &= K(x, y) = 1 && \text{if } y - x \in \langle \theta \rangle \\ &= 0 && \text{otherwise,} \end{aligned}$$

where $\langle \theta \rangle$ denotes the (countable) subgroup of S^1 generated by θ . Clearly, $K = 0$ almost surely relative to the product measure $\lambda \times \lambda$, but $K(T^i x, T^j x) = 1$ along every orbit of T . Therefore, (19) fails. It is not difficult to modify the function K so that the limit fails to exist with probability one. \square

3.2. SLLN for self-intersections. It is not much more difficult to prove the law of large numbers (1), by incorporating an additional observation from [8]. First, observe that Proposition 1 can be reformulated as a statement about weak convergence of empirical distributions:

Corollary 3. *Let (\mathcal{X}, d) be a compact metric space, μ a Borel probability measure on \mathcal{X} , and $T : \mathcal{X} \rightarrow \mathcal{X}$ an ergodic, measure-preserving transformation relative to μ . For each $x \in \mathcal{X}$ and $n \geq 1$, let*

$$(20) \quad \mu_n^x := \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \delta_{(T^i x, T^j x)}$$

be the empirical measure that puts mass $1/n^2$ at each point $(T^i x, T^j x)$. Then for μ -almost every $x \in \mathcal{X}$,

$$(21) \quad \mu_n^x \xrightarrow{w*} \mu \times \mu.$$

Consequently, if $U \subset \mathcal{X} \times \mathcal{X}$ is any Borel measurable set whose topological boundary satisfies $\mu \times \mu(\partial U) = 0$, then for μ -almost every $x \in \mathcal{X}$,

$$(22) \quad \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \mathbf{1}_U(T^i x, T^j x) = \mu \times \mu(U).$$

Proof. The assertion (21) is just an equivalent form of (19), by the definition of weak-* convergence. Given (21), the statement (22) follows by elementary analysis (see, for instance, Billingsley's "Portmanteau Theorem", [1], Theorem 2.1). \square

Proof of SLLN (1). Fix $\delta > 0$ small, and consider the representation (9) for $N(L)$ when $L = n\delta$. For $\varphi \equiv 1$, the intersection kernel $H_\delta(x, y)$ takes the form of an indicator function $H_\delta(x, y) = \mathbf{1}_U(x, y)$, where $U = U_\delta$ is the set of pairs (x, y) such that the geodesic segments of length δ with initial points x, y intersect transversally. It is easily checked (see [8]) that the boundary of this set has $\nu_L \times \nu_L$ -measure 0, so Corollary 3 implies that for almost every initial point,

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} N(\gamma[0, n\delta]) = \nu_L \times \nu_L(U_\delta).$$

By the same argument as in the proof of Lemma 1, the constant on the right is $\sim \kappa_M \delta^2$ as $\delta \rightarrow 0$. Thus, the SLLN (1) follows. \square

4. WEAK CONVERGENCE OF FLUCTUATIONS

4.1. Heuristics. We begin by using the results of sections 2.1–2.2 to give a compelling — but non-rigorous — explanation of the weak convergence asserted in Theorem 1. The Hilbert-Schmidt theorem asserts that a symmetric integral kernel in the class $L^2(\nu_L \times \nu_L)$ has an L^2 -convergent eigenfunction expansion. The intersection kernel $H_\delta(u, v)$ meets the requirements of this theorem, and so its eigenfunction expansion converges in $L^2(\nu_L \times \nu_L)$:

$$(23) \quad H_\delta(u, v) = \sum_{k=1}^{\infty} \lambda_k \psi_k(u) \psi_k(v).$$

The L^2 -convergence of the series does not, of course, imply *pointwise* convergence. Nevertheless, let's proceed formally, ignoring convergence issues: Recall (Corollary 2) that the eigenfunctions are mutually uncorrelated, and so all except the constant eigenfunction ψ_1 have mean zero relative to ν_L . Thus, the representation (9) of the intersection number $N_\varphi(n\delta)$ can be rewritten as follows:

$$(24) \quad \begin{aligned} N_\varphi(n\delta) - (n\delta)^2 \kappa_g &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n H_\delta(\gamma(i\delta), \gamma(j\delta)) - (n\delta)^2 \kappa_g \\ &= \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=2}^{\infty} \lambda_k(\delta) \psi_k(\gamma(i\delta)) \psi_k(\gamma(j\delta)) \\ &= \frac{1}{2} \sum_{k=2}^{\infty} \lambda_k(\delta) \left(\sum_{i=1}^n \psi_k(\gamma(i\delta)) \right)^2. \end{aligned}$$

If the eigenfunctions ψ_j were Hölder continuous, the central limit theorem for the geodesic flow [13] would imply that for any finite K the joint distribution of the random vector

$$(25) \quad \left(\frac{1}{\sqrt{n}} \sum_{i=1}^n \psi_k(\gamma(i\delta)) \right)_{2 \leq k \leq K}$$

converges, as $n \rightarrow \infty$, to a (possibly degenerate) K -variate Gaussian distribution centered at the origin. (The central limit theorem in [13] is stated only for the case $K = 1$, but the general case follows by standard weak convergence arguments [the ‘‘Cramer-Wold device’’, as in [1], ch. 1.] Hence, for every $K < \infty$ the distribution of the truncated sum

$$(26) \quad \frac{1}{n} \sum_{k=2}^K \lambda_k(\delta) \left(\sum_{i=1}^n \psi_k(\gamma(i\delta)) \right)^2$$

should converge, as $n \rightarrow \infty$, to that of a quadratic form in the entries of the limiting Gaussian distribution.⁴

⁴That the limit distribution has the same form as required by Definition 1 follows by the spectral theorem for symmetric matrices and elementary properties of the multivariate Gaussian distribution, as follows. Suppose that the limit distribution of the random vector (25) is mean-zero Gaussian with (possibly degenerate) covariance matrix Σ ; this distribution is the same as that of $\Sigma^{1/2}Z$, where Z is a Gaussian random vector with mean zero and identity covariance matrix. Let Λ be the diagonal matrix with diagonal entries $\lambda_j(\delta)$. Then the limit distribution of (26) is identical to that of $Z^T M Z$, where $M = \Sigma^{1/2} \Lambda \Sigma^{1/2}$. But the matrix M is symmetric, so it may be factored as $M = U^T D U$, where U is an orthogonal matrix and D is diagonal. Now if Z is mean-zero Gaussian with the identity covariance matrix, then so is UZ , since U is orthogonal. Thus, $Z^T M Z$ has the same form as in Definition 1, where θ_j are the diagonal entries of D .

There are, obviously, two problems with this argument. First, the central limit theorem requires that the functions ψ_j be Hölder continuous; but since the intersection kernels H_δ are not continuous, their eigenfunctions ψ_j will not be continuous either. This problem could be circumvented by smoothing. But second, the convergence of the infinite series in (24) and the interchange of limits requires justification. The fact that the series (23) converges in $L^2(\nu_L \times \nu_L)$ is of no use here, because for any $s, t > 0$ the joint distribution of $(\gamma(s), \gamma(t))$ is *singular* relative to $\nu_L \times \nu_L$. If the kernel $H_\delta(u, v)$ were continuous and positive semi-definite, then Mercer's theorem ([4], ch. 3) would imply *pointwise* — in fact, uniform — convergence of the series; unfortunately, neither H_δ nor its smoothed version K_δ is positive semi-definite.

The way around these difficulties, it seems, is not to use the eigenfunction expansion, but instead to approximate, in C^m -norm for suitable m , the smoothed kernel $K_\delta(x, y)$ by *elementary kernels*, that is, finite sums of the form

$$(27) \quad h(x, y) := \sum_{i=1}^J \sum_{j=1}^J a_{i,j} \varphi_i(x) \varphi_j(y),$$

where the functions φ_j are C^m , with mean zero relative to ν_L , and the matrix $(a_{i,j})$ is symmetric. A repetition of the argument given above shows that for any such kernel h ,

$$T^{-1} \iint_{[0,T]^2} h(\gamma(s), \gamma(t)) ds dt \xrightarrow{\mathcal{D}} G$$

where G is a (possibly degenerate) quadratic form in independent unit Gaussian random variables. Thus, to prove Theorem 1 it will suffice to show that the error incurred in approximating K_δ by h in the integral

$$T^{-1} \iint_{[0,T]^2} K_\delta(\gamma(s), \gamma(t)) ds dt$$

is small uniformly for $T > 1$. In the remainder of this section we shall show that such estimates reduce to a problem of mixing for the geodesic flow.

4.2. Multiple Mixing Rates for the Geodesic Flow. The proof of Theorem 1 will rely on an extension of DOLGOPYAT's theorem [5] on exponential mixing rates for the geodesic flow. For the sake of simplicity, since our only interest is in the case of the Liouville measure ν_L , we will state the extension only for this measure. Denote by \mathbb{E} expectation relative to ν_L , and by $\gamma_t = \gamma(t)$ the geodesic ray whose initial direction vector γ_0 is randomly chosen according to ν_L .

Theorem 3. *For each $K = 2, 3, \dots$, there exist constants $C = C_K < \infty$, $A = A_K > 0$, and $m = m_K \in \mathbb{N}$ such that for any mean-zero, real-valued functions F_1, F_2, \dots, F_K of class C^m on SM and all $t_1 < t_2 < \dots < t_K$,*

$$(28) \quad \left| \mathbb{E} \prod_{j=1}^K F_j(\gamma(t_j)) \right| \leq C \left(\min_{j \leq K} \exp\{-A(t_j - t_{j-1})\} \right) \prod_{j=1}^K \|F_j\|_m.$$

Here $\|F\|_m$ denotes the C^m norm of the function F , that is,

$$(29) \quad \|F\|_m := \min_{k \leq m} \sup_{x \in SM} |D_{i_1} D_{i_2} \cdots D_{i_k} F(x)|$$

where the maximum is over all mixed partial derivatives of order $\leq m$ in the directions of unit tangent vectors to SM .

Dolgopyat [5] proves the case $K = 2$. The proof of the extension stated above is a near-verbatim transcription of the proof in [5]; the only real difference is that here we must deal with Fourier-Laplace transforms in several variables. Define the K -point correlation function

$$(30) \quad \varrho(s_1, s_2, \dots, s_{K-1}) = \mathbb{E} \prod_{j=0}^{K-1} F_j(\gamma(t_j))$$

where $t_0 = 0$ and $s_j = t_j - t_{j-1}$, and $\gamma(t)$ represents (as usual) the geodesic flow. Denote by $\hat{\varrho}$ the Fourier-Laplace transform:

$$(31) \quad \hat{\varrho}(z_1, z_2, \dots, z_{K-1}) = \int_0^\infty \cdots \int_0^\infty \varrho(s_1, s_2, \dots, s_{K-1}) e^{\sum_{j=1}^{K-1} z_j s_j} ds_1 ds_2 \cdots ds_{K-1}.$$

The multivariable Paley-Wiener theorem (see, for instance, Theorem IX.13 [14]) implies that to prove Theorem 3 it suffices to prove that for constants $C < \infty$ and $A > 0$ depending only on the differentiability class m ,

$$(32) \quad \int_{-\infty}^\infty \cdots \int_{-\infty}^\infty |\hat{\varrho}(x_1 + iy_1, \dots, x_{K-1} + iy_{K-1})| dy_1 dy_2 \cdots dy_{K-1} \leq C \prod_{j=0}^{K-1} \|F_j\|_m$$

for all real $x = (x_1, x_2, \dots, x_{K-1})$ such that $0 \leq x_j \leq A$. This is accomplished by rewriting the Fourier-Laplace integral (31) in terms of Ruelle-Perron-Frobenius operators as in Proposition 3 of [5] (see also [11]). Our case differs from the case considered there only in that the double series in Proposition 3(c) in [5] must be replaced by a multiple series involving $K - 1$ Ruelle operators $\mathcal{L}_{z_i, \tau}$; but (this is the key point) the potential kernels $z_i \tau$ in these Ruelle operators are all multiples of the return time function τ . The estimate (32) then follows by using the estimates on powers of the Ruelle operators given in Corollary 2 of [5].

4.3. Approximation by Elementary Kernels. To make use of Theorem 3 we will approximate the centered kernels $K_\delta^*(x, y) := K_\delta(x, y) - \kappa_\varphi$ by *elementary kernels*, that is, kernels of the form (27):

$$h(x, y) := \sum_{i=1}^J \sum_{j=1}^J a_{i,j} \varphi_i(x) \varphi_j(y),$$

where the functions φ_j are C^m , with mean zero relative to ν_L , and the matrix $(a_{i,j})$ is symmetric. Existence of such approximations (in particular, the requirement that the functions φ_j have mean zero) depends crucially on the fact that the constant function 1 is an eigenfunction of K_δ (Property (P1) above): this guarantees that the kernel $K_\delta^*(x, y)$ is *centered*, that is,

$$(33) \quad \int K_\delta^*(x, y) \nu_L(dy) = \int K(x, y) \nu_L(dy) - \kappa_\varphi = 0 \quad \text{for every } x \in SM.$$

Lemma 7. *Let $g(x, y)$ be a centered, symmetric kernel of class C^m , where $m \geq 0$. Then for every $\varepsilon > 0$ there is an elementary kernel $h(x, y)$ such that*

$$(34) \quad \|g - h\|_m < \varepsilon.$$

Proof. This is completely standard except for the requirement that the component functions φ_j in the approximating kernel h be mean-zero. Suppose that g is well-approximated by a kernel h of the form (27) in which the component functions φ_j are *not* mean-zero. Denote by $\bar{\varphi}_j$ the mean of

φ_j relative to ν_L . Since $g(x, y)$ integrates to 0 against $\nu_L(dy)$ for every x , the inequality (34) implies that

$$\left| \sum_{i=1}^J \sum_{j=1}^J a_{i,j} \bar{\varphi}_j \varphi_i(x) \right| < \varepsilon, \quad \text{whence} \quad \left| \sum_{i=1}^J \sum_{j=1}^J a_{i,j} \bar{\varphi}_i \bar{\varphi}_j \right| < \varepsilon.$$

Consequently, $g(x, y)$ is also well-approximated by the kernel

$$h^*(x, y) := \sum_{i=1}^J \sum_{j=1}^J a_{i,j} \varphi_i^*(x) \varphi_j^*(y) \quad \text{where} \quad \varphi_j^*(x) = \varphi_j(x) - \bar{\varphi}_j.$$

□

We will need a quantitative version of the approximation (34), in which the C^m -norms of the terms $a_{i,j} \varphi_j(x) \varphi_i(y)$ are controlled by the size of the kernel $g(x, y)$. This can be done, but at the cost of a more stringent differentiability requirement on $g(x, y)$.

Lemma 8. *For every $m \geq 1$ there exists $C < \infty$ such that the following holds: For every centered, symmetric kernel $g(x, y)$ of class C^{2m} , the elementary kernel $h(x, y)$ in the approximation (34) can be chosen so that*

$$(35) \quad |a_{i,j}| \leq C \|g\|_{2m} (i+j)^{-m} \quad \text{and} \quad \|\varphi_j\|_m \leq 1.$$

Proof. First, use a smooth partition of the identity to localize, then use Fourier series approximations in the coordinate patches. The coordinate patches U_i in $SM \times SM$ can be chosen so that they are nearly isometric to cubes in \mathbb{R}^6 ; using a matching partition of 1, we obtain a decomposition

$$g = \sum_{k=1}^K g_k$$

where each g_k is supported by U_i and has C^{2m} -norm bounded by $C' \|g\|_{2m}$, with a constant C' independent of the kernel g . Each g_k may now be viewed as a C^{2m} -function on a cube in \mathbb{R}^6 , and as such can be expanded in a Fourier sine series. (This may have a nonzero constant term, but this will wash out later, since the original kernel g is centered.) Now $\sin kx$, as a function on $[0, 2\pi]$, has L^2 -norm $1/\sqrt{2}$, but has C^m -norm k^m . However, because g_k is of class C^{2m} , its inner product with any product of sines can be integrated by parts up to $2m$ times, so the Fourier coefficient of any sine product containing a factor $\sin kx$ will be bounded in magnitude by C''/k^{2m} . Thus, the resulting series approximation will satisfy (35) for suitable $C < \infty$. Finally, the resulting approximations to g can be modified so that the component functions φ_k are mean zero, by the same argument as in the proof of Lemma 7. □

4.4. L^2 Bounds via Approximation. Lemma 7 asserts that every centered, symmetric kernel $g(x, y)$ of class C^m can be arbitrarily well-approximated, in the C^m norm, by elementary kernels $h(x, y)$. This leaves the problem of determining how much of an error might be incurred in replacing g by h in the integral

$$(36) \quad T^{-1} \iint_{[0,T]^2} g(\gamma(s), \gamma(t)) ds dt.$$

It is obvious that if $\|g - h\|_m < \varepsilon$ then the error cannot exceed εT . But this isn't good enough for our purposes: we need the error to be of size $O(1)$ for large T . This is where the multiple-mixing rate provided by Theorem 3 comes in.

Lemma 9. For $m \in \mathbb{N}$ sufficiently large there exist constants $C = C_m < \infty$ such that if $g(x, y)$ is a centered, symmetric kernel of class C^{2m} then for all $T \geq 1$,

$$(37) \quad \mathbb{E} \left(\iint_{[0, T]^2} g(\gamma(s), \gamma(t)) ds dt \right)^2 \leq CT^2 \|g\|_{2m}.$$

Proof. Let $g(x, y)$ be a centered, symmetric kernel of class C^{2m} . By Lemma 7, for each $T \geq 1$ there exist an elementary kernel $h = h_T$ such that $g - h$ has C^m -norm smaller than $1/T^3$. Furthermore, by Lemma 8, the kernel $h = h_T$ can be chosen so that it is of the form

$$h(x, y) = \sum_i \sum_j a_{i,j} \varphi_i(x) \varphi_j(y)$$

where the functions φ_i are mean-zero with C^m -norms bounded by 1, and the coefficients $a_{i,j}$ satisfy

$$|a_{i,j}| \leq C' \|g\|_{2m} / (i + j)^m$$

Because $\sum \sum (i + j)^{-m} < \infty$ if $m \geq 3$, and since the C^m -norm dominates the supremum norm, it follows that

$$\left| \iint_{[0, T]^2} g(\gamma(s), \gamma(t)) ds dt - \iint_{[0, T]^2} h(\gamma(s), \gamma(t)) ds dt \right| \leq C'' T^{-1} \|g\|_{2m}.$$

Consequently, it suffices to establish the inequality (37) with g replaced by $h = h_T$. But because h is elementary,

$$\begin{aligned} & \mathbb{E} \left(\iint_{[0, T]^2} h(\gamma(s), \gamma(t)) ds dt \right)^2 \\ &= 4! \sum_{i_1} \sum_{i_2} \sum_{i_3} \sum_{i_4} a_{i_1, i_2} a_{i_3, i_4} \iint \iint \iint_{0 < s_1 < s_2 < s_3 < s_4 < T} \mathbb{E} \prod_{j=1}^4 \varphi_{i_j}(\gamma(s_j)) ds_1 ds_2 ds_3 ds_4 \end{aligned}$$

Now Theorem 3 applies: in particular, it implies that the inner expectation is bounded in magnitude by

$$C''' \min(\exp\{-A(s_2 - s_1)\}, \exp\{-A(s_4 - s_3)\}).$$

Thus, the quadruple integral is bounded by a constant multiple of T^2 . \square

4.5. Weak Convergence for Centered Kernels. The approximation results of section 4.3 together with the L^2 -bound provided by Lemma 9 together imply that the normalized integral (36) has, for large T , a distribution close to that of a Gaussian quadratic form:

Theorem 4. Let $g(x, y)$ be a centered, symmetric kernel of class C^∞ . Then as $T \rightarrow \infty$,

$$(38) \quad T^{-1} \iint_{[0, T]^2} g(\gamma(s), \gamma(t)) ds dt \xrightarrow{\mathcal{D}} F,$$

where F is a probability distribution in the weak closure of the set of Gaussian quadratic forms.

Proof. By Lemma 7, for each $m \geq 1$ and any $\varepsilon > 0$ there is an elementary function $h(x, y)$ of class C^m such that the difference $r := g - h$ has C^m -norm less than ε . Since h is elementary, it is

centered and symmetric; hence, so is the difference r . Therefore, by Lemma 9, if m is sufficiently large then for all $T > 1$,

$$\mathbb{E} \left(T^{-1} \iint_{[0,T]^2} r(\gamma(s), \gamma(t)) ds dt \right)^2 \leq C\varepsilon$$

where $C = C_m < \infty$ is a constant depending only on the differentiability class C^m . On the other hand, since h is elementary, it has the form

$$h(x, y) = \sum_{i=1}^J \sum_{j=1}^J a_{i,j} \varphi_i(x) \varphi_j(y),$$

that is, it is a finite, symmetric quadratic form in the functions φ_j . Consequently, by the same argument as in section 4.1 above, as $T \rightarrow \infty$,

$$T^{-1} \iint_{[0,T]^2} h(\gamma(s), \gamma(t)) ds dt \xrightarrow{\mathcal{D}} G$$

where G is a (possibly degenerate) quadratic form in independent unit Gaussian random variables. Since $\varepsilon > 0$ can be chosen arbitrarily small, it follows that for large T the random variable

$$T^{-1} \iint_{[0,T]^2} g(\gamma(s), \gamma(t)) ds dt$$

is arbitrarily close in L^2 to a random variable with a distribution close to that of a Gaussian quadratic form. The theorem follows. \square

Remark 4. Unfortunately, it seems, this argument gives no information about the limit distribution F other than the fact that it is a weak limit of distributions of Gaussian quadratic forms. The principal difficulty is in the use of the central limit theorem for the geodesic flow: Even if one knows, say, that the functions φ_i in the elementary kernel h are uncorrelated under ν_L , it does not follow that the random variables

$$\frac{1}{\sqrt{T}} \int_0^T \varphi_i(\gamma(s)) ds$$

are uncorrelated, nor that their Gaussian limits will be uncorrelated. Thus, it seems that it is impossible to relate the coefficients in the limiting quadratic form to the coefficients $a_{i,j}$ in the expansion of the elementary approximations.

4.6. Proof of Theorem 1. Theorem 4 applies to any centered, symmetric kernel of class C^∞ , and so in particular to the kernels K_δ defined by (14). Consequently, to complete the proof of Theorem 1 it suffices to show that the error in the inequalities (15) is small (in $L^1(\nu_L)$) compared to t , for large t . Recall (Remark 2) that the error in (15) is bounded by

$$(39) \quad \begin{aligned} N_\varphi(\gamma[-\delta, t + \delta]) - N_\varphi(\gamma[\delta, t - \delta]) &= M_\varphi(\gamma[-\delta, \delta], \gamma[-\delta, t + \delta]) \\ &\quad + M_\varphi(\gamma[-\delta, t + \delta], \gamma[t - \delta, t + \delta]) \end{aligned}$$

where $M_\varphi(\alpha, \beta)$ denotes the weighted sum of the transversal intersections between the geodesic segments α and β , that is,

$$M_\varphi(\alpha, \beta) = \sum \varphi(\theta_i)$$

where the sum is over all intersections i between α and β , and θ_i is the angle of the i th intersection. Hence, since the Liouville measure ν_L is invariant and reversible under the geodesic flow, the expected error in (15) is bounded by

$$2\mathbb{E}M_\varphi(\gamma[0, 2\delta], \gamma[2\delta, t + 4\delta]).$$

Thus, it suffices to show that this expectation is small compared to t when δ is small. This is implied by the following lemma, which completes the proof of Theorem 1.

Lemma 10.

$$(40) \quad \limsup_{\delta \rightarrow 0} \sup_{t \geq 1} t^{-1} \mathbb{E}M_\varphi(\gamma[0, \delta], \gamma[0, t]) = 0.$$

Proof. Recall that geodesic segments of length $\leq \varrho$ can intersect at most once, provided $\varrho > 0$ is sufficiently small. Hence, if $\delta < \varrho$ then $M_\varphi(\gamma[0, \delta], \gamma[0, t])$ cannot be larger than $\|\varphi\|_\infty t / \varrho$. Thus, the expectation in (39) is $O(t)$. The problem is to prove that the implied constant shrinks to 0 as $\delta \rightarrow 0$. For this we will use the SLLN (3) (proved in section 3) and an averaging trick. The averaging trick is this: since M_φ is additive,

$$(41) \quad \delta^{-1} \int_{s=0}^t M_\varphi(\gamma[s, s + \delta], \gamma[0, t]) ds \leq 2N_\varphi(\gamma[0, t]) \leq \delta^{-1} \int_{s=-\delta}^{t+\delta} M_\varphi(\gamma[s, s + \delta], \gamma[0, t]) ds.$$

This implies that $N_\varphi(\gamma[0, t])$ is bounded above by $\|\varphi\|_\infty t(t + 2\delta) / (\delta\varrho)$, and so in particular the random variables $N_\varphi(\gamma[0, t]) / t^2$ are uniformly bounded. Therefore, the SLLN implies convergence of expectations:

$$(42) \quad \lim_{t \rightarrow \infty} 2\mathbb{E}N_\varphi(\gamma[0, t]) / t^2 = \kappa_\varphi.$$

Next, by additivity of M_φ and reversibility of the geodesic flow relative to the Liouville measure,

$$\begin{aligned} \mathbb{E}M_\varphi(\gamma[s, s + \delta], \gamma[0, t]) &= \mathbb{E}M_\varphi(\gamma[0, \delta], \gamma[0, t - s - \delta]) \\ &\quad + \mathbb{E}M_\varphi(\gamma[0, \delta], \gamma[0, s + \delta]). \end{aligned}$$

Substituting this in (41) yields

$$2\delta^{-1} \int_{s=\delta}^{t-\delta} \mathbb{E}M_\varphi(\gamma[0, \delta], \gamma[0, s]) ds \leq 2\mathbb{E}N_\varphi(\gamma[0, t]).$$

Since $M_\varphi(\gamma[0, \delta], \gamma[0, s])$ is nondecreasing in s , it follows that

$$\mathbb{E}M_\varphi(\gamma[0, \delta], \gamma[0, t/2]) / t \leq 2\delta \mathbb{E}N_\varphi(\gamma[0, t]) / t(t - \delta).$$

The desired result (40) now follows from the convergence of expectations (42). \square

Remark 5. A more sophisticated argument, using the central limit theorem for the geodesic flow, shows that the errors in the inequality (15) are actually of order \sqrt{t} .

5. CENTRAL LIMIT THEOREM FOR $N_{\varphi;f}(t)$

5.1. Localized Intersection Kernels. In proving Theorem 2 there is no loss of generality in considering only *nonnegative* functions f , so we shall assume throughout that $f : M \rightarrow \mathbb{R}$ is a nonnegative, C^∞ function and that the smoothing function φ satisfies the hypotheses of Theorem 2. When convenient, we will view f as a function on SM : that is, for any $u = (x, \theta) \in SM$, set $f(u) = f(x)$. Recall that the f -localized self-intersection counts $N_{\varphi;f}(t)$ are obtained by summing $f(x_i)\varphi(\theta_i)$, where x_i and θ_i are the locations and angles of the self-intersections, and the sum is

over all self-intersections i of the geodesic segment $\gamma[0, t]$. As for the global self-intersection counts, the localized counts can be represented by *intersection kernels*. As is the case for the global intersection kernels, the obvious local kernels are not continuous, despite the fact that the functions φ and f are both C^∞ . Consequently, we define smooth kernels k_δ as in Section 2.3:

$$(43) \quad k_\delta(u, v) = \sum_i \delta^{-2} p(s_i/\delta) p(t_i/\delta) f(x_i) \varphi(\theta_i)$$

where i, s_i, t_i , and p are as in the definition (14). By the same argument as in Lemma 5, if $\delta > 0$ is sufficiently small then the sum (43) contains at most one nonzero term, and so k_δ extends to a symmetric, C^∞ function on $SM \times SM$. In addition, by the same argument as in Lemma 6,

$$(44) \quad N_{\varphi;f}^*(\gamma[\delta, t - \delta]) \leq N_{\varphi;f}(\gamma[\delta, t - \delta]) \leq N_{\varphi;f}^*(\gamma[-\delta, t + \delta])$$

where

$$(45) \quad N_{\varphi;f}^*(\gamma[a, b]) := \frac{1}{2} \iint_{[a,b]^2} k_\delta(\gamma(r), \gamma(s)) dr ds.$$

Furthermore, by the same argument as in Section 4.6, the errors in the inequalities (44) are of order $O(t)$. Since the limit relation (8) involves fluctuations of order $t^{3/2}$, it follows that the errors in (44) can be ignored in proving Theorem 2. Thus, it suffices to prove, for some small $\delta > 0$, that as $t \rightarrow \infty$,

$$(46) \quad \frac{N_{\varphi;f}^*(\gamma[0, t]) - A_{\varphi;f} t^2}{\sigma t^{3/2}} \implies \Phi.$$

Corollary 4. (SLLN) Define $2A_\delta = \kappa_\varphi \mathbb{E} f_\delta$ where the expectation is with respect to the Liouville measure. Then for all sufficiently small $\delta > 0$,

$$(47) \quad A_\delta = A_{\varphi;f} \quad \text{and} \quad \lim_{t \rightarrow \infty} \frac{N_{\varphi;f}^*}{t^2} = \lim_{t \rightarrow \infty} \frac{N_{\varphi;f}}{t^2} = A_{\varphi;f} \quad a.s.$$

Proof. Proposition 1 implies that $N_{\varphi;f}^*(t)/t^2$ converges almost surely to A_δ . But the inequalities (44) and the fact that the errors are of order $O(t)$ imply that the limit constant does not depend on δ . \square

5.2. Lead Eigenfunction. The primary difference between the global intersection kernel K_δ and the local kernel k_δ is that the constant function 1 is not, in general, an eigenfunction of the local kernel. To see this, define

$$(48) \quad f_\delta(u) := \frac{1}{\kappa_\varphi} \int_{SM} k_\delta(u, v) \nu_L(dv).$$

Lemma 11.

$$\lim_{\delta \rightarrow 0} \|f_\delta - f\|_\infty = 0.$$

Proof. Because k_δ is non-zero only for pairs u, v at distance $< 2\delta$, the value of f in (43) will be close to $f(u)$ for all small δ , uniformly for $u \in SM$. (Here we are viewing f as a function on SM .) Hence,

$$|f(u)K_\delta(u, v) - k_\delta(u, v)| \leq \max_{d(a,b) \leq \delta} |f(a) - f(b)|.$$

Since $K_\delta 1 = \kappa_\varphi$, by (P1) of section 2.3, the result follows from the continuity of f . \square

5.3. Cohomology and the Central Limit Theorem. By Lemma 11, if f is not cohomologous to a constant, then neither is f_δ provided $\delta > 0$ is sufficiently small. This follows trivially from the definition:

Definition 3. A C^∞ function $g : SM \rightarrow \mathbb{R}$ (or \mathbb{C}) is said to be a *coboundary* relative to the geodesic flow if it integrates to zero along every closed geodesic; similarly, g is said to be *cohomologous* to a constant a if $g - a$ is a coboundary.

It is not so easy to find nonconstant functions that are cohomologous to constants, but it is quite easy to construct a function g that is not cohomologous to a constant: Take two closed geodesics α and β that do not intersect⁵ on M , and let $g : M \rightarrow \mathbb{R}$ be any C^∞ , nonnegative function that is identically 1 along α but vanishes in a neighborhood of β . In fact, the existence of non-intersecting closed geodesics yields the existence of a large class of functions that are not cohomologous to constants:

Proposition 2. Let $\varepsilon > 0$ be the distance in M between two non-intersecting closed geodesics α and β . Then no C^∞ , nonnegative, function $g : M \rightarrow \mathbb{R}$ that is not identically zero and whose support has diameter less than ε is cohomologous to a constant.

Proof. By hypothesis, g vanishes on at least one of the geodesics α, β . Because closed geodesics are dense in SM , their projections are dense in M . Thus, since g is not identically 0, there is a closed geodesic ξ on which the average value of g is positive. \square

The importance of the concept of cohomology to us is its relation to the central limit theorem for the geodesic flow: If g is *not* cohomologous to a constant, then there exists⁶ $\sigma = \sigma_g > 0$ such that if $\gamma(0)$ is chosen at random from the Liouville measure, then

$$(49) \quad \frac{1}{\sigma\sqrt{t}} \left\{ \int_0^t g(\gamma(s)) ds - \int_{SM} g d\nu_L \right\} \Longrightarrow \Phi.$$

5.4. Proof of Theorem 2. Rewrite the relation (45) as

$$(50) \quad 2N_{\varphi;f}^*(\gamma[0, t]) = 2\kappa_\varphi \iint_{[0,t]^2} f_\delta(\gamma(r)) dr ds - 2A_\delta t^2 + \iint_{[0,t]^2} \tilde{k}_\delta(\gamma(r), \gamma(s)) dr ds$$

where

$$\tilde{k}_\delta(u, v) = k_\delta(u, v) - \kappa_\varphi(f_\delta(u) + f_\delta(v)) + A_\delta.$$

Corollary 4 implies that $A_\delta = A_{\varphi;f}$ for all small $\delta > 0$. By Proposition 2, the function f is not cohomologous to a constant, and so by Lemma 11 neither is f_δ , provided $\delta > 0$ is sufficiently

⁵That there are non-intersecting closed geodesics on any compact surface of *constant* negative curvature follows from the standard representation of such a surface as a geodesic polygon in the hyperbolic plane with sides identified, because non-crossing sides will be non-intersecting closed geodesics after the identifications are made. It then follows that there are non-intersecting closed geodesics on any compact surface of variable negative curvature, because all variable-negative-curvature metrics on a compact surface can be obtained by smooth deformation of the constant curvature metric, and such deformations preserve transversal intersections of closed geodesics. For more detail, see, for instance, [7].

⁶See [13] for the assertion that $\sigma_f > 0$ if and only if f is not cohomologous to a constant. The definition of cohomology used there is different from ours, but it is easily shown that the two definitions are equivalent.

small. Consequently, the fluctuations of the first integral in (50) are of order $t^{3/2}$, and the central limit theorem implies that for $\sigma = \sigma_{f,\delta} > 0$,

$$\frac{1}{\sigma t^{3/2}} \left\{ \kappa_\varphi \iint_{[0,t]^2} f_\delta(\gamma(r)) dr ds - A_{\varphi;f} t^2 \right\} \implies \Phi.$$

On the other hand, since the kernel \tilde{k}_δ is C^∞ , symmetric, and *centered*, the fluctuations of the second integral in (50) are of order t , by Theorem 4. Theorem 2 follows. \square

Acknowledgment. Thanks to JAYADEV ATHREYA for a helpful discussion.

REFERENCES

- [1] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- [2] Joan S. Birman and Caroline Series. An algorithm for simple curves on surfaces. *J. London Math. Soc. (2)*, 29(2):331–342, 1984.
- [3] Joan S. Birman and Caroline Series. Geodesics with bounded intersection number on surfaces are sparsely distributed. *Topology*, 24(2):217–225, 1985.
- [4] R. Courant and D. Hilbert. *Methods of mathematical physics. Vol. I*. Interscience Publishers, Inc., New York, N.Y., 1953.
- [5] Dmitry Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998.
- [6] Ulrich Krengel. *Ergodic theorems*, volume 6 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 1985. With a supplement by Antoine Brunel.
- [7] Steven P. Lalley. Mostow rigidity and the Bishop-Steger dichotomy for surfaces of variable negative curvature. *Duke Math. J.*, 68(2):237–269, 1992.
- [8] Steven P. Lalley. Self-intersections of closed geodesics on a negatively curved surface: statistical regularities. In *Convergence in ergodic theory and probability (Columbus, OH, 1993)*, volume 5 of *Ohio State Univ. Math. Res. Inst. Publ.*, pages 263–272. de Gruyter, Berlin, 1996.
- [9] Péter Major. Limit theorems for nonlinear functionals of Gaussian sequences. *Z. Wahrsch. Verw. Gebiete*, 57(1):129–158, 1981.
- [10] Maryam Mirzakhani. Growth of the number of simple closed geodesics on hyperbolic surfaces. *Ann. of Math. (2)*, 168(1):97–125, 2008.
- [11] Mark Pollicott. On the rate of mixing of Axiom A flows. *Invent. Math.*, 81(3):413–426, 1985.
- [12] Mark Pollicott and Richard Sharp. Angular self-intersections for closed geodesics on surfaces. *Proc. Amer. Math. Soc.*, 134(2):419–426 (electronic), 2006.
- [13] M. Ratner. The central limit theorem for geodesic flows on n -dimensional manifolds of negative curvature. *Israel J. Math.*, 16:181–197, 1973.
- [14] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975.
- [15] Robert J. Serfling. *Approximation theorems of mathematical statistics*. John Wiley & Sons Inc., New York, 1980. Wiley Series in Probability and Mathematical Statistics.
- [16] Murad S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31:287–302, 1974/75.
- [17] Harold Widom. *Lectures on integral equations*. Notes by David Drazin and Anthony J. Tromba. Van Nostrand Mathematical Studies, No. 17. Van Nostrand, 1969.
- [18] Norbert Wiener. The ergodic theorem. *Duke Math. J.*, 5(1):1–18, 1939.

UNIVERSITY OF CHICAGO, DEPARTMENT OF STATISTICS, 5734 UNIVERSITY AVENUE, CHICAGO IL 60637

E-mail address: lalley@galton.uchicago.edu